DOCUMENT RESUME

ED 102 206 TM 004 316

AUTHOR Osborn, William C.
TITLE Process Versus Product Measures in Performance Testing. Professional Paper No. 16-74.
INSTITUTION Human Resources Research Organization, Alexandria, Va.
REPORT NO HumRRO-PP-16-74
PUB DATE Oct 73
NOTE 9p.; Paper presented at the Military Testing Association Conference (San Antonio, Texas, October 1973)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Academic Achievement; *Formative Evaluation; Military Training; *Performance Tests; Skill Analysis; *Summative Evaluation; Task Performance; *Training

ABSTRACT

Performance tests are used in training evaluation to (1) certify student achievement, and (2) diagnose weaknesses in the instructional system. Proficiency measures that focus on task outcomes (product) normally provide data relevant to the first purpose, whereas measures of how the tasks are carried out (process) pertain to the second. Time or cost factors sometimes preclude the use of product measures, leaving measures of task process as the only available criteria for evaluating training outcomes. Instances in which process measures are typically substituted for product measures are described in this paper with reference to the types of tasks for which the substitution is valid and those for which it is invalid. Theoretical and practical issues pertaining to the use and misuse of process measures are discussed. (Author)

Professional
Paper
16-74
HumRRO PP-16-74

HumRRO

Process Versus Product Measures in Performance Testing

William C. Osborn

Presented at the Military Testing
Association Conference
San Antonio, Texas October 1973

HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street • Alexandria, Virginia 22314

October 1974

The Human Resources Research Organization (HumRRO) is a nonprofit corporation established in 1969 to conduct research in the field of training and education. It is a continuation of The George Washington University Human Resources Research Office. HumRRO's general purpose is to improve human performance, particularly in organizational settings, through behavioral and social science research, development, and consultation.

Published
October 1974
by
HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street
Alexandria, Virginia 22314

PREFATORY NOTE

This paper explores why process measures are widely substituted for measures of task product and suggests ways to assess task outcome in a relevant fashion. The paper was presented at the Military Testing Association Conference, San Antonio, Texas, in October 1973.

Mr. Osborn, former Senior Staff Scientist for the Human Resources Research Organization, Division No. 2, Fort Knox, Kentucky, has recently been appointed Director of Division No. 2.

PROCESS VERSUS PRODUCT MEASURES IN PERFORMANCE TESTING

William C. Osborn

Consider the following situations.

After training in tank gunnery, a soldier's proficiency is tested on a gunnery range. During this test he will fire several main gun rounds at targets varying in size, shape, and distance. In each case his score is determined by whether he hits the target within some specified time limit. He is certified a tank gunner if he scores above some minimum level required for qualification.

Under other circumstances, a soldier with similar training may be evaluated differently. Let's assume that ammunition is scarce or that adequate range facilities are not available. Here the soldier might have to be tested under dry-firing conditions. He would be required to take actual or miniaturized versions of targets under fire. A tester would assess in each case whether the gunner (a) acquired the target with smooth manipulation of the hand controller, (b) correctly ranged on the target, (c) achieved the proper sight picture, (d) squeezed the firing switch without losing the sight picture, and (e) fired within some allotted time. The gunner is qualified if he performed each of the five procedural steps correctly on some minimum number of targets.

In the first situation described, a task outcome or product measure—target hits—is the basis for evaluating gunners, whereas in the second instance, correct task procedure or a process measure is the basis for evaluation. Although somewhat oversimplified, the contrasting approaches to performance testing drawn in these two examples illustrate the focus of this paper.

I am chiefly interested in the use of performance tests to evaluate the results of training. Let me introduce this topic by summarizing what the training evaluator considers to be the ideal use of product and process measures. Performance tests are used in training evaluation to serve two purposes: (a) to certify student achievement, and (b) to diagnose weaknesses in the instructional system. In the use of such tests, proficiency measures which focus on task outcomes (products) normally provide data relevant to the first purpose, whereas measures of how the tasks are carried out (process) pertain to the second. For example, the number of targets hit by the tanker trainee would be the product measure by which his qualification as a tank gunner is assessed. However, if he fails to qualify we would also like to know why—where was his training weak? This is where process measures are useful. If the gunner consistently missed targets, was it because he ranged incorrectly, obtained an improper sight picture, or was not able to maintain the gun lay during firing? This type of data is useful in diagnosing areas of training deficiency, and is essential in efficiently remediating trainees.

Both product and process measures are important—even critical—when used for their respective purposes in evaluating the results of training.

Product and Process as Measures of Student Achievement

With this background, I would like to focus on the use of these types of performance measures in certifying student achievement. In testing a student to determine whether he is qualified to advance to the next level of training, or

ultimately out of training and on to the job, we usually prefer to use a product score. Before a man is certified as a gunner, we would like to have him demonstrate that he can hit targets; before certification as a navigator, he should actually demonstrate that he can get from point A to point B.

Although it may safely be said that every task has a purpose, many performance tests are used which employ process measures only in evaluating student achievement or job readiness. Why is this the case? Is the substitution of process for product measurement justified? If so, when? If not, how may the test developer improve his methods? These are the questions I will address.

Types of Tasks

Before exploring in more detail the issue of why process measures are so widely substituted for measures of task product, it will first be helpful to consider three types of tasks:

(1) Tasks in which the product is the process.
(2) Tasks in which the product always follows from the process.
(3) Tasks in which the product sometimes follows from the process.

Relatively few tasks are of the first type—those in which product and process are the same. These are normally tasks which serve an aesthetic purpose, such as gymnastic exercises or springboard diving. Close order drill is a good military example. Here we see that the outcome or product of the task is no more or less than the correct execution of steps in task performance—that is, the process.

More tasks are of the second type mentioned—those in which the product invariably follows from the process. Fixed-procedure tasks typically fall in this category. Troubleshooting an electrical circuit, disassembling a rifle, and implanting a land mine are examples. In tasks of this type the procedural steps are known, observable, and comprise the necessary and sufficient conditions for task outcome; so, if process is correctly executed, task product necessarily follows.

A great many job tasks are of the third type where the product is less than fully conditional on the process. In other words, with these types of tasks the process may appear to have been correctly carried out but the goal or product was not achieved. This can happen for one of two reasons: either (a) we are unable to fully specify the necessary and sufficient steps in task performance or (b) we cannot or do not accurately measure them.

In aim-firing a rifle, for example, we are interested in knowing whether a soldier is standing with his body properly oriented to the target, face properly positioned on stock, rifle sling in correct position, lead arm perpendicular, and firing arm parallel to ground—whether he is breathing correctly, has a good sight picture, and squeezes the trigger. Presumably, if this process is followed the rifleman will hit the target. Assuming that we have identified all essential steps in rifle firing, and, further, that we can reliably measure their correct execution, then the task is of the second and not the third type described above.

However, in practice, we know that our best efforts to evaluate execution of this particular task are not sufficient to warrant substituting process for product measurement. In other words, sometimes the target is missed even though in the judgment of a skilled evaluator the rifleman did everything right. Therefore, either because we are not absolutely certain that we have identified all necessary steps in the firing process or because we cannot accurately assess the execution of some of them, we ultimately qualify a rifleman on the basis of whether he hits the target.

In reflecting on the nature of these three types of tasks, an important implication emerges regarding the role of product measurement in testing task performance: Because of the interchangeability of process and products for tasks of the first two

types, it does not really matter which measure is used to assess proficiency; but for tasks of Type 3, product measurement is very important. Nevertheless, performance tests for many of the latter type of tasks do not attempt to measure product. Why is this so?

Problems in Product Measurement

The reasons largely stem from practical considerations in which the measurement of task product is viewed as either too costly, too dangerous, or, for other reasons, simply too impractical. In testing such performances as hand-to-hand combat, for example, where task product would take the form of disabling a hostile enemy, the test developer is normally limited to requiring the demonstration of task process. Similarly, in a first aid task like controlling the bleeding from an external wound, the person tested is, for obvious reasons, asked only to demonstrate the process. Or, in removing a jammed round from a weapon, it is considered impractical to actually jam a round in order to create a valid test situation, so again only the steps in task performance are measured.

Many examples can be found in the area of interpersonal behavior. In instructor training, for instance, the military instructor trainee is traditionally evaluated by having him prepare and deliver a block of instruction during which he is judged on such process factors as: "stood erect," "had good eye contact with audience," "could be heard in the back of the room," "used visual aids effectively," and "covered all points in the lesson plan." Although clearly the product of instruction is student learning, I believe it is seldom, if ever, used as the criterion for qualifying an instructor trainee—probably because it would involve a time consuming and impractical method of evaluation. A similar example is that of a recruiter's task in delivering a persuasive speech to a student audience. If the product of this task could be measured, it would be in terms of the number in the audience who later contacted the recruiter with an interest in enlisting. But, again, because of its implausibility as a measure of student achievement, product gives way to process and the recruiter trainee's persuasive speech is evaluated in much the same way as was described for the instructor trainee.

Dealing With Problems of Product Measurement

Those of you involved in performance testing can think of many more instances in which product measurement is not used. Certainly, some of these are justified by cost or safety considerations—but others are not. I believe that test developers often fail to see the importance of measuring task outcome; or perhaps they merely slight the importance when faced with practical limitations in its measurement. Whatever the motivation, I believe they do not strive hard enough to overcome resource problems which hinder measurement of task product, and give in too easily to the simplistic approach of measuring task process.

The question that a test designer should ask himself in this situation is: *If I use only a process measure to test a man's achievement on a task, how certain can I be that he will also be able to affect the product or outcome of the task?* Where the degree of certainty is substantially less than that to be expected from normal measurement error, the test designer should pause and reconsider ways in which time and resource limitations can be compromised in achieving at least an approximation to product measurement. Although there will remain instances in which product measurement simply cannot be achieved, we will discover many others where, through some imaginative thinking, we can devise simulations enabling us to assess task outcome in a more relevant fashion.

In testing the student instructor, for instance, I see no compelling reason why we should not get away from the "charm school" approach to evaluation. Why not simply have him conduct a brief instructional session for a small group of students (perhaps his peers), with his achievement measured in terms of whether his students have accomplished the instructional objective? In the case of the recruiter trainee's speech, evaluating task product is more difficult; but surely a measure closer to task outcome could be achieved—perhaps a student panel representing the potential audience could be employed to view and rate the appeal of videotaped trainee speeches.

In evaluating critical motor skills, such as those involved in extracting a jammed round from a weapon or in controlling bleeding from a wound, it would seem that relatively low-cost simulators could be devised for use in testing task outcome. Hand-to-hand combat very likely represents a case in which ultimate task product simply cannot be measured. However, in a similar vein, the Army is now experimenting with an intriguing method of assessing the outcome of an infantry squad combat exercise. The principal feature of the method entails each participant having a number printed on his helmet and an inexpensive scope mounted on his rifle; then, during the exercise a soldier may "kill" by correctly reporting an enemy's number, or "be killed" by allowing his number to be sighted by the enemy. Number size and scope power have been carefully calibrated from empirical data so that the probability of a simulated "kill" is highly correlated with the expected outcome in actual battle. This is an excellent example of an innovative method of achieving product measurement on a task that heretofore had been subject to process evaluation.

Obviously, from these examples we can see that the accomplishment of product measurement is not always a simple matter; but it is a demanding and essential goal to be pursued by the performance test developer if his products are to be relevant to real-world behavior.

BIBLIOGRAPHIC DATA SHEET	1. Report No. HumRRO-PP-16-74	2.	3. Recipient's Accession No.
4. Title and Subtitle PROCESS VERSUS PRODUCT MEASURES IN PERFORMANCE TESTING			5. Report Date October 1974 6.
7. Author(s) William C. Osborn			8. Performing Organization Rept. No. PP-16-74
9. Performing Organization Name and Address Human Resources Research Organization (HumRRO) 300 North Washington Street Alexandria, Virginia 22314			10. Project Task/Work Unit No. 11. Contract/Grant No.
12. Sponsoring Organization Name and Address Human Resources Research Organization (HumRRO) 300 North Washington Street Alexandria, Virginia 22314			13. Type of Report & Period Covered Professional Paper 14.

15. Supplementary Notes

Research performed at HumRRO Division No. 2, Fort Knox, Kentucky; presented at the Military Testing Association Conference, San Antonio, Texas, October 1973.

16. Abstracts

Performance tests are used in training evaluation to (a) certify student achievement, and (b) diagnose weaknesses in the instructional system. Proficiency measures that focus on task outcomes (product) normally provide data relevant to the first purpose, whereas measures of how the tasks are carried out (process) pertain to the second. Time or cost factors sometimes preclude the use of product measures, leaving measures of task process as the only available criteria for evaluating training outcomes. Instances in which process measures are typically substituted for product measures are described in this paper with reference to the types of tasks for which the substitution is valid and those for which it is invalid. Theoretical and practical issues pertaining to the use and misuse of process measures are discussed.

17. Key Words and Document Analysis. 17a. Descriptors

Infantry training
Instructors
Military training
Performance evaluation
Training devices

17b. Identifiers. Open-Ended Terms

Performance tests
Product measures
Process measures
Training evaluation

17c. COSATI Field Group 0509

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
Distribution of this document is unlimited.	UNCLASSIFIED	7
	20. Security Class (This Page) UNCLASSIFIED	22. Price

FORM NTIS-35 (10-70)

USCOMM-DC 40329-P71